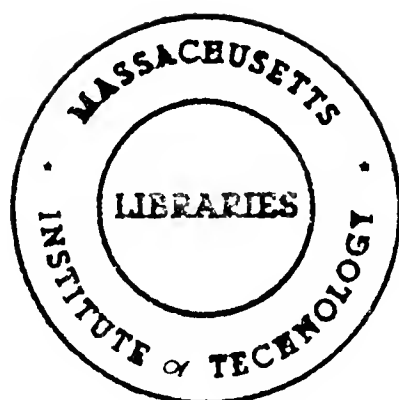




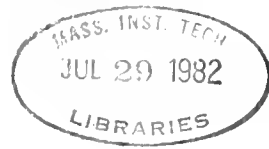
BASEMENT



HD28
.M414

No 1309-82

Dewey



A LOGIT MODEL OF BRAND CHOICE
CALIBRATED ON SCANNER DATA

Peter M. Guadagni*

John D. C. Little**

A LOGIT MODEL OF BRAND CHOICE
CALIBRATED ON SCANNER DATA

Peter M. Guadagni*

John D. C. Little**

May 1982

SSM Working Paper: # 1304-82

*Management Decision Systems, Inc., 200 Fifth Avenue, Waltham, MA 02254

**Massachusetts Institute of Technology, E53-350, Cambridge, MA 02139

W.L.F. L. 1000
JUL 2 1961
RECEIVED

ABSTRACT

Optical scanning of the Universal Product Code in supermarkets provides a new level of detail and completeness in household panel data and makes possible the construction of more comprehensive brand choice models than hitherto possible. A multinomial logit model calibrated on 32 weeks of purchases of regular ground coffee by 100 households shows high statistical significance for the explanatory variables of brand loyalty, size loyalty, presence/absence of store promotion, regular shelf price and promotional price cut. The model is quite parsimonious in that the coefficients of these variables are modeled to be the same for all coffee brand-sizes. Considering its parsimony, the calibrated model predicts surprisingly well the share of purchases by brand-size in a hold-out sample of 100 households over the 32 week calibration period and a subsequent 20 week forecast period. Discrepancies in prediction are conjectured to be due in part to missing variables.

Three short term market response measures are calculated from the model: regular shelf price elasticity of share, percent share increase from a promotion with a median price cut, and promotional price cut elasticity of share. Responsiveness varies across brand-sizes in a systematic way with large share brand-sizes showing less responsiveness in percentage terms. On the basis of the model a quantitative picture emerges of groups of loyal customers who are relatively insensitive to marketing actions and a pool of switchers who are quite responsive.

1. INTRODUCTION

The automatic recording of purchases at the point of sale has opened new opportunities for building and testing models of customer behavior. In particular, data now being collected by optical scanning of the Universal Product Code (UPC) in supermarkets should permit a careful examination of various customer choice models developed in recent years. We shall here use scanner collected panel data on coffee purchases to calibrate a multinomial logit choice model and examine both its scientific quality as a representation of consumer behavior and its potential usefulness for marketing decision-making.

Scanner data usually comes in two forms: Store data and panel data. Store data provides individual item (UPC) sales and price by store by week. In addition, the companies supplying data may collect information on other store activities such as special display, coupon redemption, retail advertising, and shelf-space allocation. Panel data presents histories of purchases for a sample of households. A cooperating household displays an identification card at checkout. The store clerk keys the household number into the cash register, thereby causing the purchase record to be segregated and stored. Over time this creates a longitudinal customer history.

Scanner data has special advantages. It directly records the sales act of the individual customer at the item level, thereby avoiding the pipeline effects found in factory shipments and warehouse withdrawals. Initial cost is low since the data is a spinoff from the store's basic transaction process. However, subsequent processing and storage can be expensive. Scanner data is potentially deliverable to users with great speed, although how fast it will actually be delivered will depend on its time-value in decision-making. The data tends to be very accurate because it is part of

the stores cash collection and financial accounting process. However, at present some purchases enter the cash register around the scanner. Store procedures determine the percentage of goods scanned and, as more stores use the data for inventory control and other purposes, accuracy will further increase. Compared, to the quality control problems of diary panels, scanner measurement is remarkably unobtrusive, accurate and complete across products and therefore represents a significant step forward in information about products sold primarily through supermarkets.

The greatest advantage of scanner data, however, is that it provides the competitive environment of the customer decision. Conventional diary data tells what the customer bought and its price but does not identify the other products, prices, and marketing activities impinging on the customer at the time of purchase. Similarly, standard warehouse withdrawal or store audit data for a geographic region does not describe the competitive situation within individual stores in the way scanner store data does. It is this rich, disaggregated detail that offers hope for new levels of consumer and market understanding. Just as Galileo's telescope stimulated astronomy and new measurements have driven advances in theory throughout the history of science, we can hope for progress here.

Current scanner data services fall in three broad categories: (1) groups of stores in single markets (2) national samples of stores and (3) instrumented markets. A group of stores in a single market offers opportunities for convenient in-store experiments using scanner-collected store data. When panels of store-loyal shoppers are added, the experimentation possibilities increase. Coupons, for example, can be validly tested for the first time, and models of individual customer behavior with respect to price and promotion become possible.

National store samples permit manufacturers to make generalizations that might be distrusted if made from a single market. Good random samples are still difficult or impossible to obtain because of the irregular geographic distribution of scanners but this will change. Empirical studies of store data can determine sales response to promotion and price at the store level and, by looking across stores, information can be gathered on retailer's response to manufacturers' promotional offerings.

The most dramatic new service stimulated by scanners, however, is the instrumented market. This is a small to medium sized city with scanners in all major grocery stores. In addition, the city is pre-selected for high cable television coverage and split cable hardware is introduced so that household panels can be set up to receive different television advertisements. In-store observations can also be conducted, and so a remarkably complete picture of the shopper's marketing environment is possible: sales, prices, promotions, advertising (both newspaper and television), coupons, display and shelf-facings. The setting is ideal for a variety of testing: new products, advertising, store promotions, etc. The instrumented market has two seeming drawbacks. One is the standard test market issue of projection to a national response. The other is an efficiency question; whereas most syndicated data services sell the same data to all manufacturers, the instrumented markets are being sold with each category exclusive to one customer. Such an arrangement protects new product and other tests. The result, however, is that costs are spread over a small fraction of the possible customers.

Much measurement and model building lies ahead using these various services. The present paper is one step. It will focus on a choice model applicable to household data collected in panels associated with individual

stores or instrumented markets, supplemented by store data on the shopping environment. To indicate the task at hand, Figure 1 displays an example of scanner collected data. The market share of a major brand of coffee in a panel of 100 households shows great variation over a year's time both in overall trend and in specific peaks. We wish to understand and predict such behavior.

2. THE MULTINOMIAL LOGIT CHOICE MODEL

The multinomial logit model computes the probability of choosing an alternative as a function of the attributes of all the alternatives available. The model has the appeal of being stochastic and yet admitting decision variables. Various authors have employed it in marketing. Hlavac and Little (1966) use a modified version to represent the probability that an automobile buyer purchases a car at a particular dealership. Silk and Urban (1978) imbed the model in their pre-test-market evaluation process for new products. Punj and Staelin (1978) employ the logit to describe students' choice of business schools. Gensch and Recker (1979) provide a general exposition and compare the fitting ability of the logit to that of regression for shoppers' choosing grocery stores. The logit has an even more extensive history of application in the field of transportation planning, particularly for predicting an individual's choice of mode of travel, e.g., car or bus (Domencich and McFadden, 1975).

2.1 Axiomatic view

As a choice model, the multinomial logit permits an axiomatic derivation which we briefly outline. Consider an individual, i , confronted with a choice from a set, S_i , of alternatives. In our setting the alternatives will be different products in a category. We suppose that:

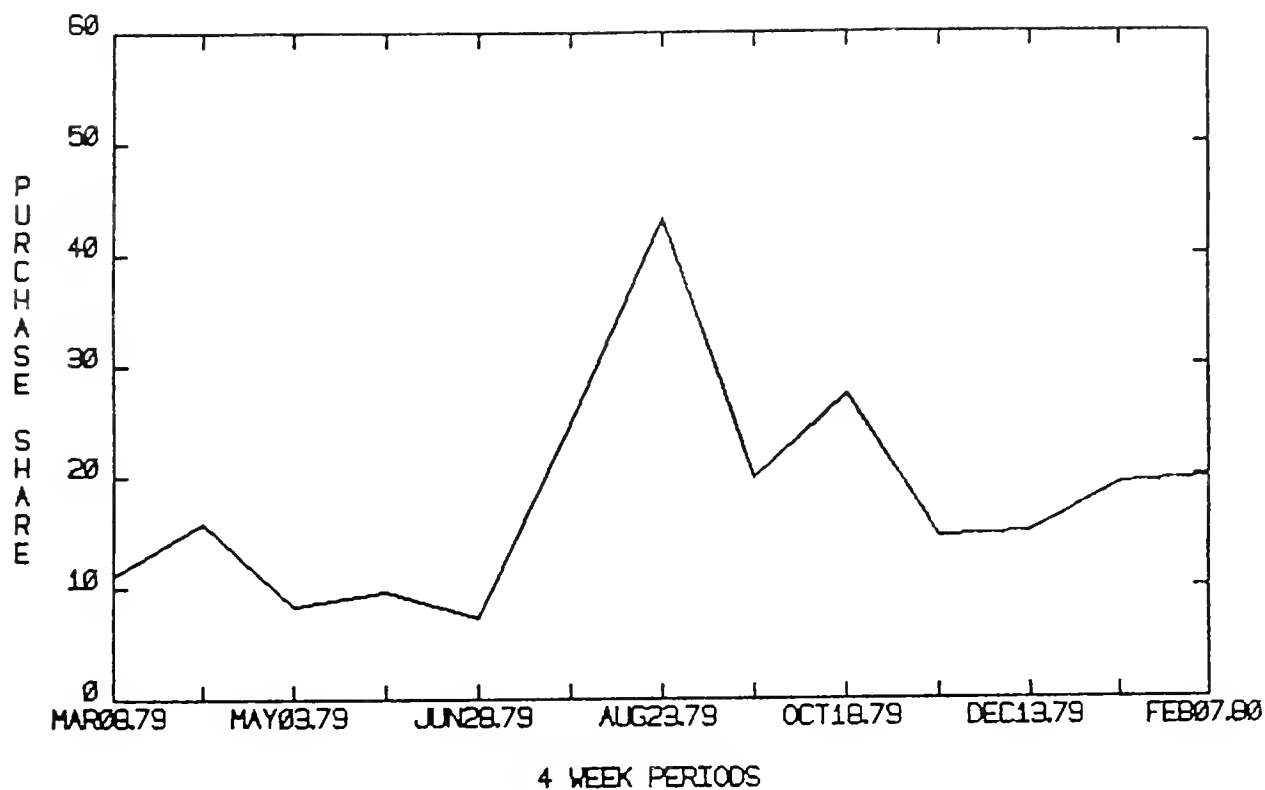


Fig. 1 : The share of purchases of a major coffee brand-size as recorded in a panel of 100 Kansas City households shows great variability. (Dates shown are the starting days of four week periods.)

(1) Alternative $k \in S_i$ holds for the individual a preference or utility,

$$u_k = v_k + \epsilon_k, \quad (1)$$

where v_k = a deterministic component of i 's utility, to be calculated from observed variables, and

ϵ_k = a random component of i 's utility, varying from choice occasion to choice occasion, possibly as a result of unobserved variables.

(2) Confronted by the set of alternatives, individual, i , chooses the one with the highest utility on the occasion. I.e., the probability of choosing k is

$$p_k = P \{u_k \geq u_j, j \in S_i\} \quad (2)$$

(3) The ϵ_k , $k \in S_i$, are independently, identically distributed random variables with a Weibull (Gnedenko, extreme value) distribution

$$P(\epsilon_k \leq \epsilon) = e^{-e^{-\epsilon}} \quad -\infty < \epsilon < \infty \quad (3)$$

This form of the distribution appears to fix the mean and variance of ϵ quite arbitrarily, since (3) has a mean .575 and a variance 1.622, both dimensionless. A more general form would include a further location parameter and a scale parameter. However, any location parameter, even one dependent on k , can be absorbed into v_k without loss of generality and, since the scaling of utility is arbitrary, we can set it so that the variance of the ϵ_k is the 1.622 value implied by (3). Notice, however, that this procedure produces larger utility values in a model that explains more variance than in one that explains a lesser amount. We shall observe this phenomena in our empirical work.

Given assumptions (1) - (3), it can be shown (Theil, 1969; McFadden, 1974) that the individual i 's choice probabilities have the remarkably

simple form

$$p_k = e^{v_k} / \sum_{j \in S_i} e^{v_j} \quad (4)$$

This expression is known as the multinomial logit.

We note two properties that we shall refer back to later. First, since (4) can be written

$$p_k = 1 / \sum_j e^{(v_j - v_k)},$$

it follows that utility is undetermined to the extent of an additive constant. Thus, for example, if a price variable has an inflationary trend that adds a constant to all alternatives, p_k will not be affected.

Second, p_k is S-shaped in v_k when other v_j are held constant. Therefore, as shown in Figure 2, very large or very small values of v_k make p_k flat and insensitive to changes in v_k .

In our case the individual choice-makers are households. We do not know whether their behavior satisfies the assumptions used to derive (4). However, the concept of utility (or preference or attractiveness) as a latent variable and a choice probability that is some normalized function of that variable has much appeal and a long history (Luce 1959, McFadden 1981).

A chief complaint about (4) is that, if we add to the alternatives a new one essentially identical to some existing alternative, say the k^{th} , the new alternative might reasonably be expected to split k 's probability and leave the others untouched, but, by (4), will instead reduce the probabilities of all alternatives. The issue is whether choices satisfy the assumption of "independence from irrelevant alternatives." Various schemes have been proposed for overcoming the potential difficulty (McFadden 1981). Many of them involve arranging the alternatives into a hierarchy or tree

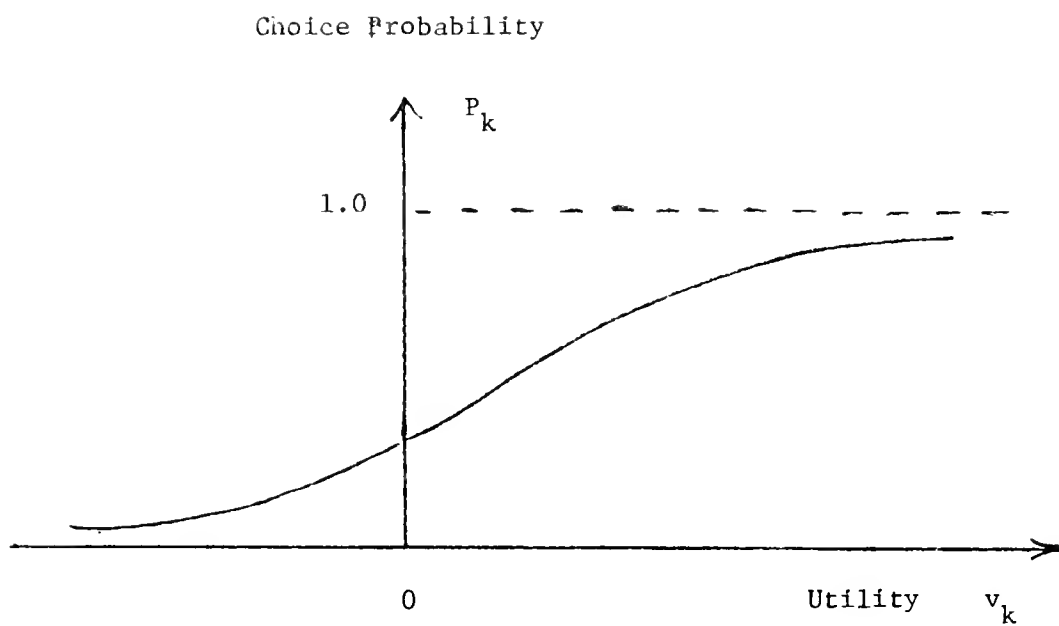


Figure 2. Choice probability is S-shaped in utility v_k .

structure that groups similar alternatives. Thus a tree representation of coffees might have all instant brands in one branch and all regular brands in another.

We shall stay with the simple multinomial (4) adopting a pragmatic view. Tests of the calibrated model will determine whether it works well. In setting up the application, however, we shall endeavor to avoid its known pitfalls.

2.2 Linear Utility

The deterministic component of a customer's utility for alternative k will be expressed as a linear function of observed variables, called the attributes of k . Some of these will be attributes of the product (e.g., price) and others may be attributes of the customer or the environment (e.g., income or store) that differentially favor one alternative over another for some reason. In general

$$v_k^i = \sum_{j \in T} b_{jk} x_{jk}^i \quad (5)$$

where x_{jk}^i = observed value of attribute j of alternative k for customer i ,

b_{jk} = utility weight of attribute j of alternative k .

We shall drop the superscript i when it is not required for clarity.

From a modeling point of view it is convenient to break the attributes into two classes.

(1) $T_k = \{\text{attributes unique to alternative } k\}$. A product may have features that other products do not have but which contribute to its utility. For such attributes $j \in T_k$, the coefficients may be denoted b_{jk} and variables by x_{jk}^i .

(2) $T_C = \{\text{attributes common to all alternatives}\}$. Although an attribute such as price might be assigned uniquely to each product, a more

parsimonious model would use price as a single attribute across all products. In this case price will have the same coefficient for all alternatives. For such attributes $j \in T_C$, the coefficients may be denoted b_j and variables x_{jk}^i .

In specialized form (5) becomes

$$v_k^i = \sum_{j \in T_k} b_{jk} x_{jk}^i + \sum_{j \in T_C} b_j x_{jk}^i \quad (6)$$

Although a linear form for utility is a natural place to start model-building, we note that linearity for v_k still leaves the choice probability quite nonlinear in the observational variables, x_{jk} . The numerator of p_k can be written

$$e^{v_k} = \prod_{j \in T} e^{b_{jk} x_{jk}}$$

so that the model is, in a sense, more multiplicative than additive.

2.3 Calibration

Equations (4) and (6) present the model. In practice we cannot observe either utilities or probabilities directly. Rather, we observe choices and attribute values. The data consist of a set of choice records for each individual:

$$y_k^i(n) = \begin{cases} 1 & \text{if customer } i \text{ chooses alternative } k \text{ on the } n^{\text{th}} \text{ choice occasion} \\ 0 & \text{otherwise,} \end{cases}$$

along with the values of the attributes of the alternatives on each choice occasion:

$$x_{jk}^i(n) = \begin{cases} \text{value of attribute } j \text{ for product } k \text{ on } n^{\text{th}} \text{ purchase occasion} \\ \text{for customer } i. \end{cases}$$

In our case a choice or observation is the purchase of a product by a customer on the occasion of buying within the product class. The b_{jk} and

b_j are unknown constants to be determined by calibration. For calibration purposes, each attribute is thought of as having a complete set of data across alternatives for each observation, even though an attribute unique to a specific alternative does not appear in the utility expressions for other alternatives. To handle this situation, such an attribute is assigned a zero value for alternatives to which it is not relevant.

Calibration is done by maximum likelihood, using (4) and (6) to calculate the likelihood function. The program used is that of Manski and Ben Akiva (Ben-Akiva, 1973). McFadden (1974) has shown that the maximum likelihood parameter estimates are consistent, asymptotically efficient and normally distributed under very general conditions.

2.4 Quality of Fit

Measures of quality of fit and parameter estimation guide model specification and help appraise the success of the calibration.

(1) t-values for coefficients.

The Manski-Ben Akiva program generates t-statistics for each coefficient in the calibrated model.

(2) U^2 for model.

Probabilistic models pose special difficulties in overall evaluation. Whereas regression models offer residuals and R^2 as ready indicators of fit, a logit model predicts only probabilities which must then be compared to actual choices. Hauser (1978) presents a set of statistics based on information theory which are useful for probabilistic models. One of these is U^2 , equivalent here to the likelihood index ρ^2 .

U^2 can be described as the proportion of uncertainty associated with a designated null model that is explained by the model being tested. The statistic is somewhat analagous to R^2 in that they both have a range of 0

to 1 and indicate degree of variability explained. U^2 tends, however, to have a lower value for an excellent fit. U^2 equals zero if the tested model predicts probabilities identical to the null model and unity if predictions are perfect (i.e., the predicted probability is 1.0 and correct). Any null model can be used. For simplicity we consistently take the case of all alternatives equally likely.

U^2 can be calculated from the output of a maximum likelihood estimation program by

$$U^2 = 1 - L(x)/L_0 \quad (7)$$

where $L(x) = \log$ likelihood of tested model with explanatory variables, x , and L_0 is the log likelihood of the null model.

(3) Chi-squared tests of model significance.

If one model, say A, can be formulated as a restriction (subset) of the parameters of the tested model, say B, then $L = 2 \log$ [likelihood ratio of model B to model A] is χ^2 distributed with degrees of freedom equal to the difference in degrees of freedom between model B and model A. See Hauser (1978). This test helps determine whether adding a parameter or set of parameters is worthwhile.

(4) Aggregate share.

For a given population the average probability of choosing an alternative is the expected share of choices for that alternative. We can compare actual with expected share. Such comparisons plotted over time offer valuable visual representations of quality of fit.

3. APPLICATION TO PACKAGED GOODS

Market share is an aggregation of individual customer choices. If we can understand how and why households choose one product over another, we

shall gain insight into the reasons for a product's success or failure. Scanner panel data and the logit model offer an opportunity for increasing such knowledge about packaged goods.

In applying the model, the alternatives will be products, but their exact level of aggregation and which ones to include in the relevant set are not necessarily obvious. Should different flavors, or colors, or sizes be treated as different products or lumped together? When product choice is hierarchical we can induce similarity by defining alternatives at the same level in the hierarchy. For example, within the category of vegetables, the choice might be among fresh, frozen and canned. However, in studying brands, we might focus on competition within, say, canned peas. In the case of coffee, a natural grouping exists between instant and ground forms and another between caffeinated and decaffeinated products.

Another crucial issue is the degree of homogeneity in the customer population. We would like to consider customers identical in the sense that one set of utility weights applies to all. This will not imply that all customers behave in the same manner because the attribute values will vary from customer to customer.

For example, households are well known to be heterogeneous with respect to purchase probabilities because of differing brand preferences or loyalties. To model this we can introduce a customer loyalty attribute defined as a weighted sequence of past purchases of a brand. To make the model parsimonious, we can make loyalty a common attribute across brands, i.e., use the same b_j coefficient for all.

Customers also frequently show a loyalty to a particular package (cans vs. bottles in soft drinks) or size (one pound vs. two pounds or three pounds in coffee). Such preferences can be introduced as customer

attributes in the same way as brand loyalty. By bringing in past customer purchase behavior as an explanatory variable through loyalty attributes, we conveniently model purchase probability heterogeneity while treating customers homogeneously.

4. REGULAR GROUND COFFEE

Coffee is a frequently purchased product actively marketed by manufacturers and retailer alike and so makes an excellent subject for building and testing a choice model. Price changes are relatively common because of fluctuations in commodity markets. Coffee is also a traditional supermarket loss leader with frequent promotions during which the store takes one or more coordinated marketing actions such as reducing price, putting up a special display or running an advertisement in the local media.

4.1 Data

Our data base consists of ground coffee store and panel data from four Kansas City supermarkets for the 78 week period September 14, 1978 to March 12, 1980. The data has been collected by Selling Areas-Marketing, Inc. (SAMI) and kindly made available for this research.

The store sales data contains weekly item movement for each UPC in the category as well as the shelf price for each item each week. The customer panel has about 2000 households each of which has indicated it makes 90% or more of its purchases at one of the stores collecting the data. A single purchase record contains the household number, the date of purchase, the UPC, and the price paid. The panel has been cleaned to eliminate households with reporting gaps that indicate lack of store loyalty and also to omit households joining the panel in the middle of the period. The resulting static sample has been further reduced to exclude non-users of ground

coffee, these being defined as households making less than five purchases during the total time period. Many families remain, however, and, of them, 200 have been chosen at random. A group of 100 forms the calibration sample and the second 100 a hold-out sample, reserved for testing the final model. The calibration group made 1037 purchases in the chosen brand-sizes during the 32 week period, March 8, 1979 to October 17, 1979 used for calibration. Sixteen of the purchases had missing store data leaving 1021 usable purchases. The holdout group made 921 purchases in the corresponding 32 week period plus another 666 in a 20 week post period October 18, 1979 to March 5, 1980 used for projection. Five of these had incomplete data leaving 1582 purchases for evaluative purposes.

In addition to the information generated by the store computers, we have local newspaper advertisements that assist in identifying promotional activity.

4.2 Alternatives

Our selection of product alternatives builds on the research of Urban, Johnson and Brudnick (1981) about the structure of the coffee market. They find that customers tend to choose first between ground and instant coffee and then between caffeinated and decaffeinated products. To obtain a relatively homogeneous set of alternatives we restrict consideration to regular (caffeinated) ground coffee.

Another issue is brand vs. size. In the Kansas City market, the popular sizes are one pound and three pound which we shall call "small" and "large". (Included in "small" are certain 13 and 14 ounce sizes advertised as producing the same number of cups as a pound.) We found no evidence to suggest that customer choice was hierarchal on either brand or size. However, different sizes of the same brand are clearly different products

from both retailer's and customer's point of view. Customers show distinct size loyalty and retailers promote sizes separately. Therefore we model brand-sizes.

Ground coffees come as regular, drip and automatic. However, given the brand-size, all grinds are priced the same and promoted together. A household's choice depends primarily on its coffee making equipment. We therefore combine UPC's across grinds to construct the brand-size alternatives.

Two low share brand-sizes (less than 1% of the purchases of the calibration sample) were dropped for lack of observations, leaving as the set of alternatives the 8 largest selling brand-sizes.

Sizes		Brands				
Small	A	B	C	D	E	
Large	A	B	C			

4.3 Attribute Variables

The latent customer utility will be expressed as a linear function of attribute variables consisting of one unique attribute for each alternative and a set of attributes common across all alternatives. Each attribute variable will have a value for each alternative for each purchase, for example, the n^{th} purchase of customer i .

(1) Unique to alternative. The utility function for a brand-size will include an additive constant specific to that alternative. This is accomplished by a set of dummy variables:

$$x_{0k}^i(n) = \begin{cases} 1 & \text{for alternative } k \\ 0 & \text{otherwise} \end{cases}$$

Each alternative, k , has its own variable except that one of them, say the N^{th} , must be omitted to avoid a singularity in the maximum likelihood estimation, since, as previously remarked, the utilities are undetermined to the extent of an additive constant. The value of b_{0N} can be taken as zero. The resulting brand-size specific constants, b_{0k} , serve to capture any uniqueness an alternative has that is not captured by other explanatory variables, insofar as describing average choice behavior over all observations is concerned. One consequence of introducing the b_{0k} is that the average predicted probability of an alternative will equal its share of purchases over the observations. No other alternative-specific variables will be used.

(2) Common across alternatives. The major control variables we shall consider for the retailer are regular shelf price, the presence or absence of a promotion for the brand-size and, if there is a promotion, the amount of price-cut (possibly zero) during the promotion. The first of these attribute variables is price:

$$x_{1k}^i(n) = \text{regular shelf price of brand-size } k \text{ at time of customer } i\text{'s } n^{\text{th}} \text{ coffee purchase. (dol/oz.)}$$

If brand-size k is on promotion at the purchase occasion, its regular shelf price is its price prior to the start of the promotion. The second variable is promotion:

$$x_{2k}^i(n) = \begin{cases} 1 & \text{if brand-size } k \text{ was on promotion at time of customer } i\text{'s } n^{\text{th}} \text{ coffee purchase} \\ 0 & \text{otherwise} \end{cases}$$

Unfortunately, no data on feature or display activity is available for the stores in our sample. We do, however, have information on three events that would be expected to accompany a promotion: (1) short term (one to three weeks) price reduction, (2) mention of the item in the store's weekly

newspaper advertisement, and (3) an unusually high short-term movement for the product, chosen as sales greater than two standard deviations above the average for the product without promotion. Since any one of these events alone may be a faulty indicator, we infer a promotion only if two out of three of these events takes place in the given week. Notice that this means that the unusual movement indicator, which is potentially circular as a purchase explainer, never defines a promotion by itself.

The 0-1 promotion variable will miss some of the variation among promotions, such as the difference in response between an end-aisle display and a small shelf-talker. However, we can explain some of the variation by adding a depth of discount or promotional price cut variable to the model:

$$x_{3k}^i(n) = \begin{cases} \text{promotional price cut on brand-size } k \text{ at time of customer } i\text{'s } n^{\text{th}} \text{ coffee purchase (dol/oz.).} \end{cases}$$

The variable is zero when there is no promotion of k .

An important marketing question is whether a promotional purchase is, in some sense, as good as an ordinary purchase. For example, will a household that switches to a brand when it is on promotion be as likely to repurchase the brand at a later time as would be the case for a non-promotional purchase? Shoemaker and Shoaf (1977), Dodson, Tybout and Sternthal (1978) and Jones and Zufryden (1981) have found that a promotional purchase decrease the likelihood of a subsequent purchase of that brand. We can look for this effect by introducing lagged promotion variables.

$$x_{4k}^i(n) = \begin{cases} 1 & \text{if customer } i\text{'s previous purchase of coffee was a promotional purchase of an alternative with the same brand as brand-size } k \\ 0 & \text{otherwise} \end{cases}$$

$$x_{5k}^i(n) = \begin{cases} 1 & \text{if customer } i\text{'s second previous purchase of coffee was a promotional purchase of an alternative with the same brand as brand-size } k \\ 0 & \text{otherwise} \end{cases}$$

A further set of variables depend on characteristics of the customer. As observed earlier much of the heterogeneity in purchase probability over a population, be it called preference, habit or loyalty, can be captured by observing past behavior. We choose the word loyalty to describe the customers tendency to repurchase the same brand size. Let

$x_{6k}^i(n)$ = brand loyalty for brand of brand-size k for n^{th} coffee purchase of customer i.

We define

$$x_{6k}^i(n) = \alpha_b x_{6k}^i(n-1) + (1-\alpha_b) \begin{cases} 1 & \text{if customer i bought brand of alternative k at purchase occasion (n-1).} \\ 0 & \text{otherwise} \end{cases}$$

Operationally, therefore, loyalty is taken to be the exponentially weighted average of past purchases of the brand, treated as 0-1 variables. The carry-over constant is α_b . Brand loyalty can be started up by taking $x_{6k}^i(1)$ to be α_b if the brand of alternative k was the first purchase in the data history of customer i, otherwise $(1-\alpha_b)/(\text{number of brands} - 1)$. The sum of loyalties across brands always equals 1 for a customer.

Size loyalty is analogous.

$$x_{7k}^i(n) = \text{size loyalty for the size of brand-size k for } n^{th} \text{ coffee purchase of customer i.}$$

$$= \alpha_s x_{7k}^i(n-1) + (1-\alpha_s) \begin{cases} 1 & \text{if customer i bought size of alternative k at coffee purchase (n-1)} \\ 0 & \text{otherwise} \end{cases}$$

where α_s is the carry-over constant for size. Initialization is analogous. The sum of loyalties over sizes is also 1.

For all variables except the alternative-specific dummies, the coefficients, b_j , are the same for all brand-sizes. Thus in calculating customer i's utility the same coefficients will be used for every brand-size

with respect to attributes such as price, promotion, loyalty, etc. The attribute variable values will, of course, change but not the coefficients. This makes the model remarkably parsimonious. Using the same coefficients across brand-size does not imply the same response to control variables. Response will depend on the whole marketing environment and on the customers' loyalty variables.

4.4 Calibration

The calibration database contains 100 households with 1021 coffee purchases over 32 weeks spread across eight brand-sizes. In addition, 718 purchases over the previous 25 weeks have been used to initialize the loyalty variables. Each purchase will be treated as an observation so that we are combining cross-section and time-series data. This makes the loyalty variables particularly important since they carry not only much of the cross-sectional heterogeneity but also a good part of the purchase-to-purchase dynamics.

An immediate question, however, is how to pick the smoothing constants α_b and α_s for constructing the loyalty variables. The answer is: bootstrap. Approximate α 's are used and the model developed with all useful explanatory attributes ($\alpha_b = \alpha_s = .75$). Following this, first brand then size loyalty variables are replaced by ten dummy variables, each denoting whether or not the brand (or size) was purchased on the n^{th} prior occasion ($n = 1, \dots, 10$). The relative coefficient sizes indicate the impact of the n^{th} prior purchase. The results, normalized to make the first coefficient unity, appear in Figure 3. The solid line shows the exponential decay selected. The fit is satisfactory since the number of points more than a standard deviation from the line is close to what would be expected (4 actual vs. 6 expected). The corresponding α_b and α_s determine the final loyalty variables.

Brand

Size

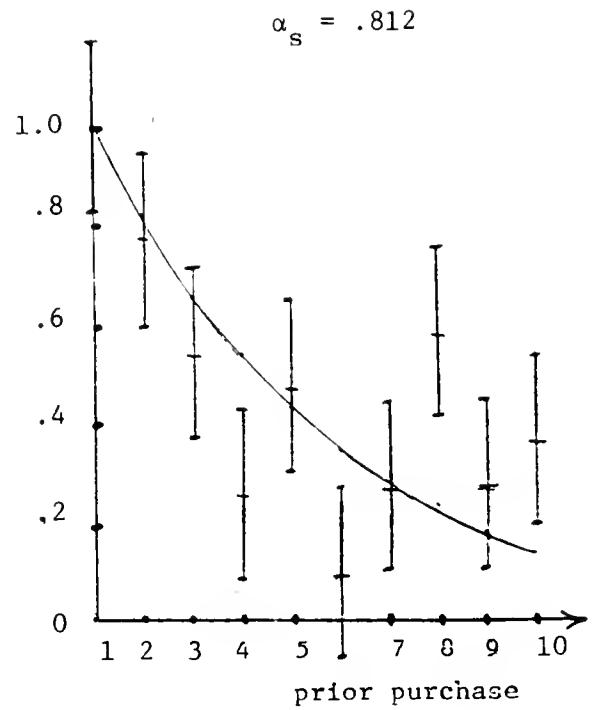
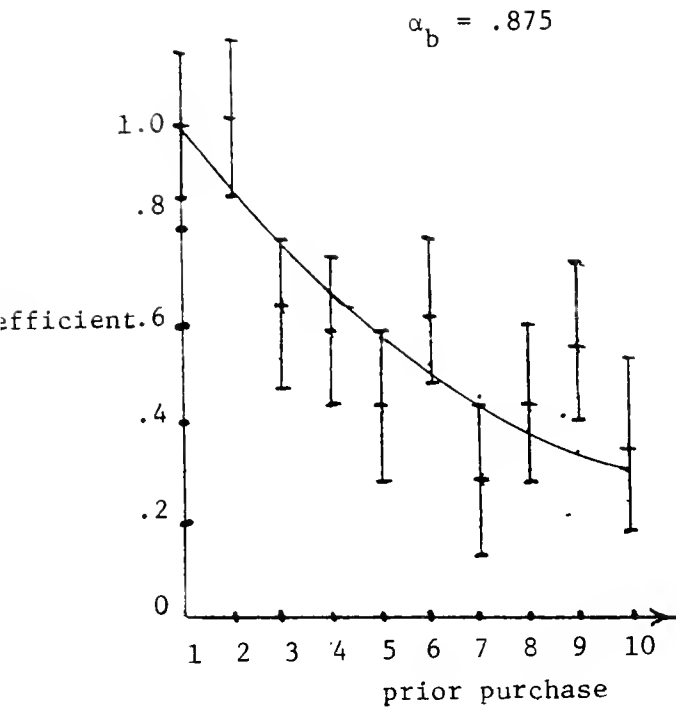


Figure 3. The smoothing parameters in the loyalty functions are fit to coefficients generated by lagged purchase variables. Bars indicate \pm one standard error.

As may be seen the decay rates for brand and size loyalty are comparable. The brand coefficients decay slightly more slowly (higher α) suggesting somewhat more loyalty to brand than size, a conclusion to be corroborated later.

5. CALIBRATION RESULTS

The calibrated logit model fits the coffee data well and provides information on a number of marketing issues. Table 1 shows the coefficients.

5.1 Quality of fit

To indicate the relative contribution of various attributes and to investigate the stability of the coefficients against changes in model specifications, the final model has been built up a few variables at a time. We can follow the changes in U^2 , the amount of uncertainty explained by the model.

The specification S1 contains only the brand-size dummy variables. The effect of this is to make each household's purchase probability for a brand-size the same and equal to that brand-size's share of total purchases. S1 produces $U^2 = .10$ relative to our reference of equally likely alternatives. In S2 the addition of the brand and size loyalty variables produces a large jump to $U^2 = .32$, demonstrating, as expected, that the loyalty variables explain a great deal of purchase behavior across households. Notice the high t-statistics. Specification S3 introduces the promotion variables with another big jump to $U^2 = .51$. Bringing in shelf price in S3 increases U^2 only a little to .52. Clearly shelf price is not moving share around the way promotion does. However, a chi-squared test of S4 relative to S3 shows statistical significance at the .99 level. When the prior promotional purchase variables are added in S5, U^2 increases to

Specification	S1	S2	S3	S4	S5	S6
U^2	.10	.32	.51	.52	.53	.54
Brand loyalty		2.78 (22.1)	3.47 (21.9)	3.47 (21.5)	3.79 (21.3)	3.92 (21.6)
Size loyalty		2.12 (14.6)	2.74 (16.1)	2.72 (15.9)	2.74 (15.9)	2.97 (15.9)
Promotion			2.22 (15.5)	2.00 (13.6)	2.07 (13.9)	2.11 (14.1)
Promotional price cut			18.12 (5.5)	29.66 (7.2)	29.20 (7.1)	29.21 (7.1)
Shelf price				-26.36 (-6.0)	-26.49 (-5.9)	-29.94 (-6.6)
Prior promotional purchase					-.60 (-3.5)	-.22 (-1.3)
Second prior promotional purchase					-.72 (-3.9)	-.46 (-2.47)
Brand-Size Constants:						
Small A	.28 (2.2)	-.12 (-.9)	-.06 (-.4)	.03 (.2)	.01 (.1)	-.09 (-.6)
Large A	0 -	0 -	0 -	0 -	0 -	0 -
Small B	.48 (3.9)	.28 (2.0)	.11 (.7)	.19 (1.2)	.20 (1.2)	.10 (.6)
Large B	-.06 (-.4)	.24 (1.6)	.05 (.3)	-.06 (-.4)	-.05 (-.3)	-.08 (-.5)
Small C	.99 (8.8)	.14 (1.0)	.31 (2.0)	.57 (3.6)	.62 (3.8)	.41 (2.5)
Large C	.38 (3.1)	.01 (.1)	.03 (.2)	.22 (1.3)	.27 (1.6)	.12 (.7)
Small D	-1.25 (-6.1)	-1.23 (-5.3)	-.66 (-2.6)	.15 (.5)	.10 (.3)	-.05 (-.2)
Small E	-1.83 (-6.8)	-1.52 (-5.2)	-.76 (-2.3)	-1.5 (-4.1)	-1.47 (-4.1)	-1.72 (-4.7)

Table 1. Calibration of coffee model for 6 specifications with increasing number of variables. Table entry shows attribute coefficient with t-statistic beneath in parentheses.

.53. The last step is to revise α_b and α_s as described earlier and rerun with the new loyalty variables. The result is S6, which we take as the final specification. It has $U^2 = .54$ indicating improved loyalty functions as a result of the bootstrapping. Although the prior promotional purchase variable is not significant in a statistical sense, we retain it because the second prior is. (We have checked the third prior and it is not.)

Notice the growth of the brand loyalty coefficient as more uncertainty is explained. This happens because the utility scaling increases to hold the residual variance of ϵ_k constant in (1), as discussed in Section 2.1.

The brand-size constants would be zero if we had found variables that explained all the differences among brands and sizes. Obviously this would be difficult to achieve and we have not, but notice that, in going from S1 to S6, most of the brand-size constants and their t-values decline.

Except for the brand-size constants, the coefficients, once introduced, tend to be rather stable throughout the various specifications, a healthy sign indicating that collinearity does not seem to be a serious issue. The only exception is the increase in the promotional price cut coefficient when regular shelf price is introduced. The increase seems reasonable, however, since the regular price is the reference point from which the price cut is measured and its absence could hinder the explanatory ability of the cut.

Visual measures of fit quality are always helpful. A direct comparison of the probability of purchase calculated by the model and the actual purchase outcomes is of little value since the latter is either 0 or 1. However, the expected number of purchases in a time period is easily calculated from the model and, when divided by total purchases, gives expected purchase share and can be compared to the actual share.

Figure 4 does this, showing predicted vs. actual purchase shares for the 8 brand-sizes over the calibration period. The curves are gratifyingly close.

As discussed earlier, the multinomial logit model implicitly assumes "independence from irrelevant alternatives" (IIA). McFadden, Train and Tye (1977) have devised a residuals test to evaluate whether the IIA assumption holds in a given case. The underlying idea is that violation of the assumption will cause systematic errors in predicted choice probabilities. The observations are divided into cells in a systematic way and a goodness of fit test applied to each brand-size. In our case none of the eight tests shows a systematic error at a the .05 level of significance. We conclude that violations of the IIA assumption are not a serious problem in our model. Details of the test appear in the appendix.

5.2 Discussion of Coefficients

All the coefficients have the algebraic signs that would be expected.

The relative importance of attributes in explaining market behavior is of considerable interest. The coefficient magnitudes per se are not too instructive because their units vary. Two better indicators are the contribution to U^2 and the t-statistic. The magnitude of the t-statistic offers an indication of explanatory importance because the numerator is the coefficient itself and the denominator is its standard error, which will tend to decrease if the data for the attribute has large variance. Such variance enhances the precision of measurement and opens the possibility of explaining considerable behavior.

Using these indicators we find brand and size loyalty most important. Brand has a larger coefficient and t-value than size, as well as a larger

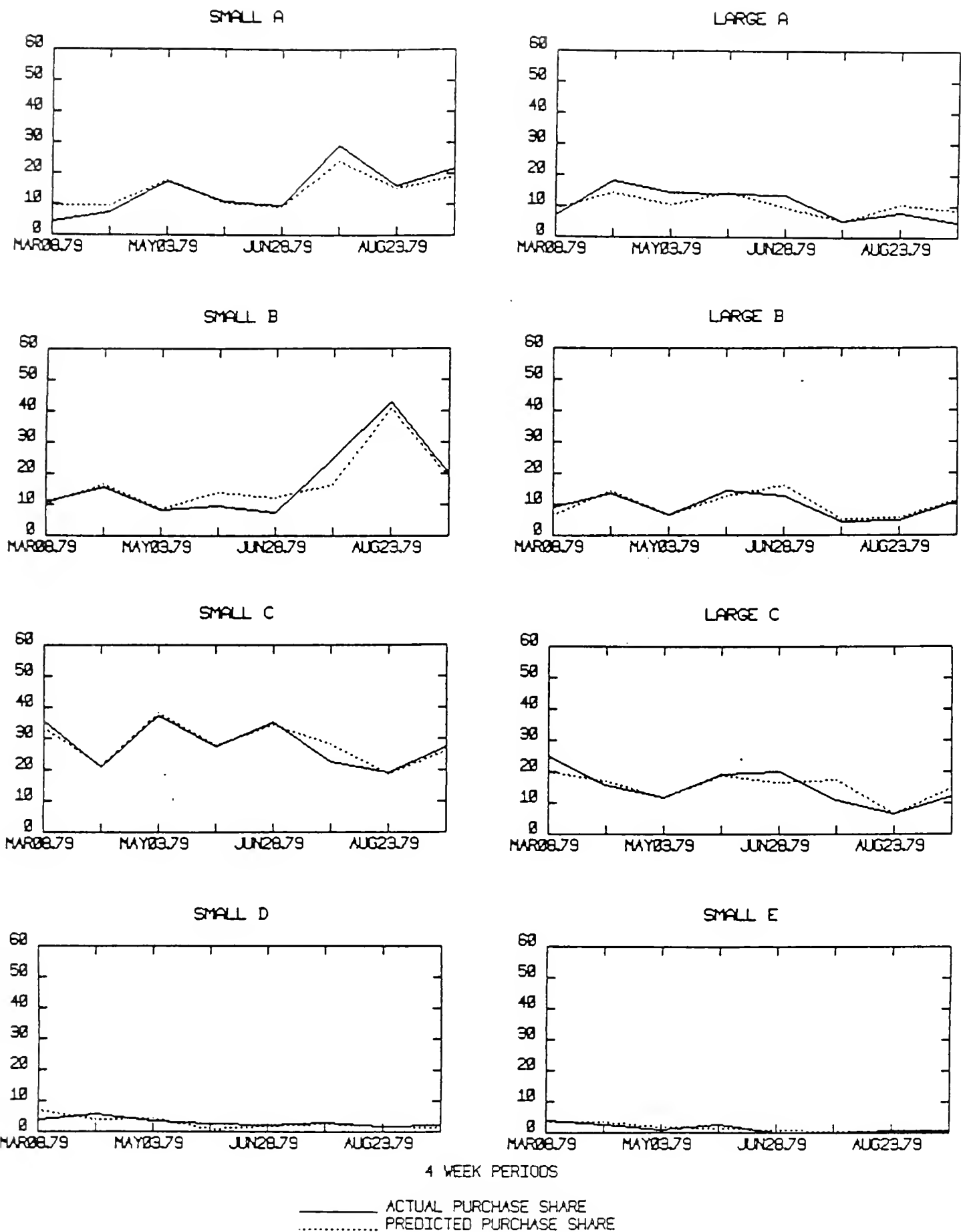


Figure 4 : Predicted share of purchases tracks actual share closely for the calibration sample over the calibration period.

carry-over constant as noted earlier. However, size loyalty is extremely important, probably much more than manufacturers and retailers realize.

The next most important attribute is the 0-1 promotion variable. Then a drop takes place to the two price variables, regular shelf price and promotional price cut, and finally another drop to the prior promotional purchase.

The brand-size constants form a special group. They can represent unique product qualities and/or specification errors. As already mentioned, if the other explanatory variables are doing a nearly perfect job, these constants should be near zero. Large A has been taken as the reference point and is zero by definition. Most of the others are small. However, it is revealing to compare brand-size constants with shares:

	Small C	Small B	Large C	Small A	Large A	Large B	Small D	Small E
Share of purchases (5):	28.4	17.1	15.5	14.0	10.6	9.8	3.0	1.6
Brand-size constant:	.41	.10	.12	-.09	0	-.08	-.05	-1.72

Perhaps a modest relationship exists, although only Small C and Small E at the extremes have significant t-values (2.5 and -4.7). One would conjecture that an explanatory variable may be missing. A top candidate would be quality of display during a promotion. Small E stands out even further than this would be expected to explain and, in fact, is a slightly different product with certain unique qualities.

Looking at the control variables further, we see that regular and promotional price coefficients have about the same magnitude. We do not expect this à priori because the promotional price is likely to be advertised in the newspapers and by signs within the store. Similarity of

coefficients does not necessarily imply similarity of price response, however, because of the other terms in the utility function.

Promotion is the most interesting control variable. Not only is its effect large but usually it is accompanied by a price cut that enhances total effect. Some marketers believe the effect of promotions to be entirely due to price but that view is not supported by our results. Even with a price cut of zero, a large effect is calculated. This is consistent with much experience that special displays (e.g., end-aisle or free standing positions) play a major role in customer response.

The past promotion variable is directed at an important question in package goods marketing: Is a customer who buys a product on promotion less likely to repurchase that product than a customer who buys it under ordinary circumstances? The answer is needed for assessing the true value of promotional activity. The work of Shoemaker and Shoaf (1977), Dodson, Tybout, and Sternthal (1978) and Jones and Zufryden (1981) indicates that the promotional purchase has a negative effect on subsequent purchase probability. The present model permits an explicit assessment of the effect. A promotional purchase affects subsequent buying through (1) brand loyalty, (2) size loyalty and (3) the prior promotional purchase variables. Suppose that a customer buys a particular brand-size on promotion. Then the contribution of that purchase to the next period's utility as a result of the loyalty variables is (using coefficients from S6 of Table 1, the α 's of Figure 3, and the definitions of x_6 and x_7):

$$(3.92) (.125) (1) + (2.97) (.188) (1) = 1.05$$

However, this contribution is diminished by the prior promotional purchase variable:

$$-(.22) (1) = -.22.$$

This leaves a net plus of .83 relative to a situation of not having purchased the brand-size at all. Carrying the process to the next purchase, the contribution of the original purchase to utility through the loyalty variables is:

$$(3.92) (.875) (.125) (1) + (2.97) (.812) (.188) (1) = .88$$

but is diminished by the second prior promotional purchase variable

$$- (.46) (1) = -.46$$

for a net of .42. The negative effect of the prior promotional purchase variables is in agreement with previous authors, but net contribution over the two periods remains positive. In other words the calculation says that the customer is more likely to purchase the promoted brand-size than if it had not been purchased in the first place, but considerably less likely than if it had been a non-promotional purchase.

6. TESTING

The holdout sample of 100 households (1609 purchases) provides an opportunity to challenge the calibrated model. Tracking share of purchases by 4-week period will be our primary method of evaluation. Not only does share loom important in practice but it varies dramatically over time and over brand-size. We ask several questions:

- (1) How well does the model predict brand-size shares in the hold-out sample during the same time period used in the calibration?
- (2) How well does the model predict share during time periods subsequent to the calibration?
- (3) Can the model, using constants derived from data that mixes together all stores, predict shares within individual stores?

In comparing predicted vs. actual, we need a measure of random variation, especially for low share brand-sizes, whose number of purchases

may be very few in a particular four-weeks. Accordingly, we calculate a standard error of predicted share. For observation i and any given brand-size, the model predicts a probability of purchase p_i . Given the null hypothesis that the model is correct, the actual purchase is binomially distributed. Letting s denote the predicted share and n the number of observations,

$$s = \sum_{i=1}^n p_i / n$$

$$SE(s) = \left[\sum_{i=1}^n p_i (1-p_i) \right]^{1/2} / n$$

where $SE(s)$ is the standard error of share.

Figure 5 displays an example of hold-out tracking, the case of Small A. The solid line in the middle is the model prediction, the upper and lower solid lines are ± 1.64 SE and so form approximately a 90% confidence band. The loyalty variables are initialized in a pre-period September 14, 1978 through March 7, 1979. Tracking then takes place during the calibration period from March 8, 1979 through October 17, 1979 and on into the post period October 18, 1979 through March 5, 1980.

As may be seen, agreement is remarkably good. A general upward trend is captured. So is a promotional peak and then a downward trend that is interrupted temporarily by another promotion. The model is impressive in tracking these ups and downs. We credit this to its detailed information about the store environment at the time purchase (namely, price, promotion, and promotional price cut for all competitors) plus information about the customer's purchase history summarized in the brand and size loyalty variables.

Of special interest is the good quality of tracking in the period October 18 to March 5, since here we not only are using a hold-out sample

share of
purchases

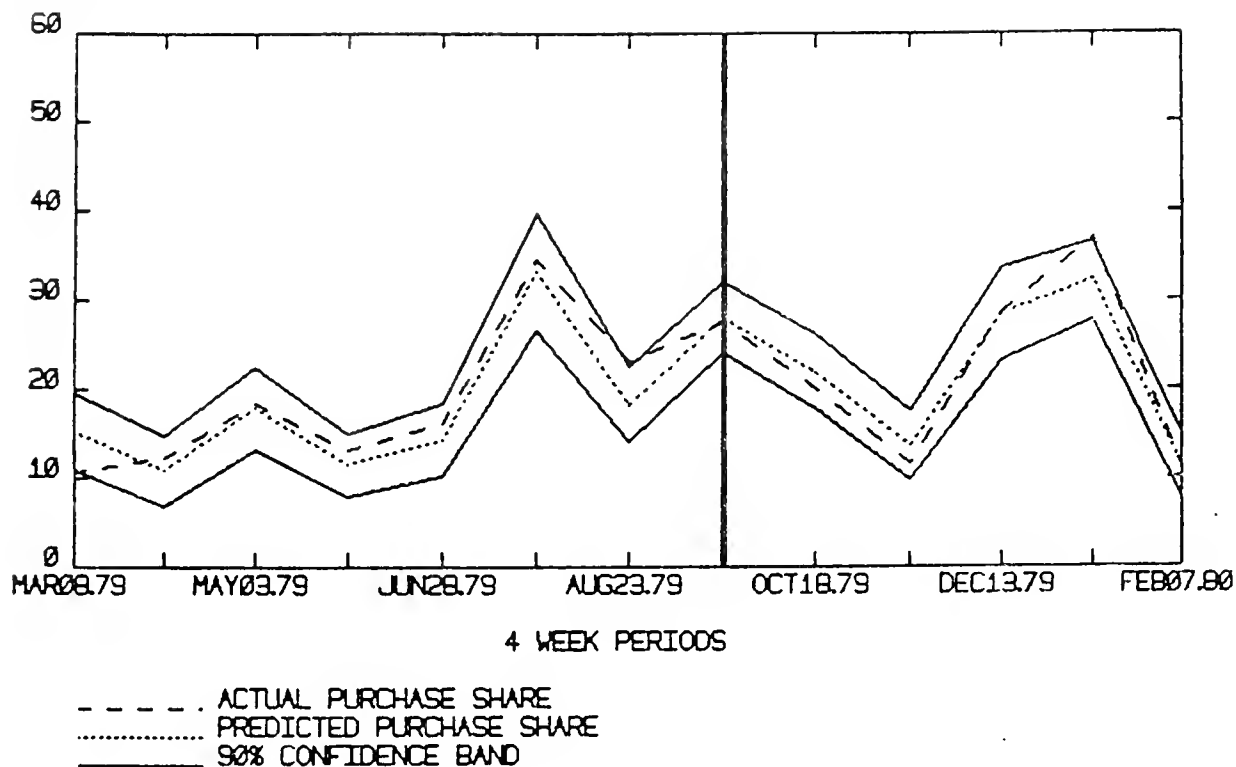


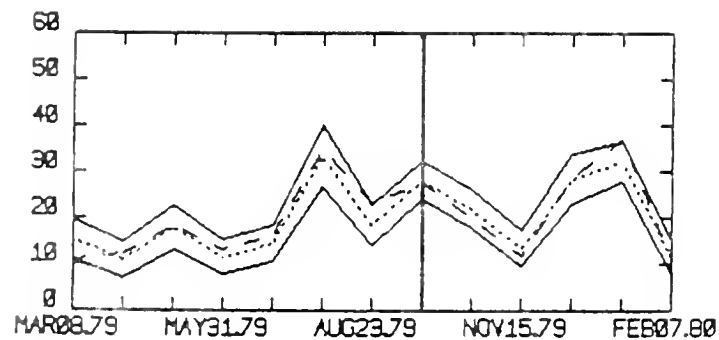
Fig. 5. Tracking of Small A purchase share in a holdout sample of 100 customers shows actual share falling almost entirely within the confidence band. The vertical line separates the calibration time period from the forecast period. During the forecast period loyalty variables are constructed from simulated purchases.

but also are forecasting outside the time period used for calibration. We are, of course, using the actual environmental variables during this period. A further point arises, however. Two of the explanatory attributes driving predicted share are loyalty variables. Loyalty uses past sales in its construction. Only past sales are used and prediction is of the probability of purchase on the next sale. Still the presence of the loyalty variables means that some of the sales data in the forecast period could, in a certain sense, enter the forecast.

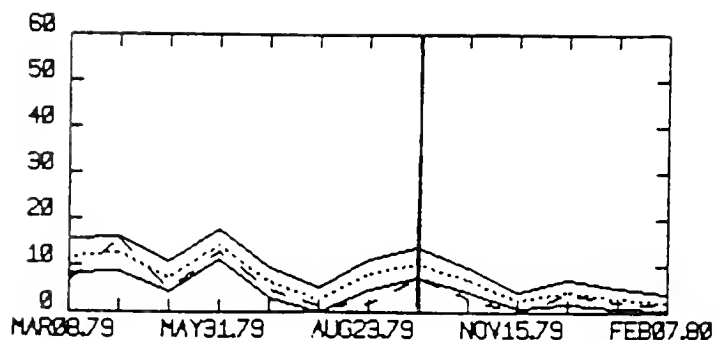
To circumvent this, we have devised the following procedure: after October 18, each time the probabilities of brand-size choice are calculated for a purchase occasion, a random number is drawn to select the brand-size to be designated as purchased, this being done so that the prior probability for a brand-size conforms to its model-calculated value. As the forecast period proceeds, a synthetic purchase history unfolds for each customer. These simulated purchases are used to update the loyalty variables at each stage. We see that, nevertheless, tracking is good in the post-calibration period.

Figure 6 displays corresponding plots for all brand-sizes. We see that, even in the severe test of a completely new set of customers, the model follows the turns and trends of share quite well. Enough points lie outside the 90% confidence band, however, to suggest the model has not captured all phenomena and, indeed, we know some variables are missing.

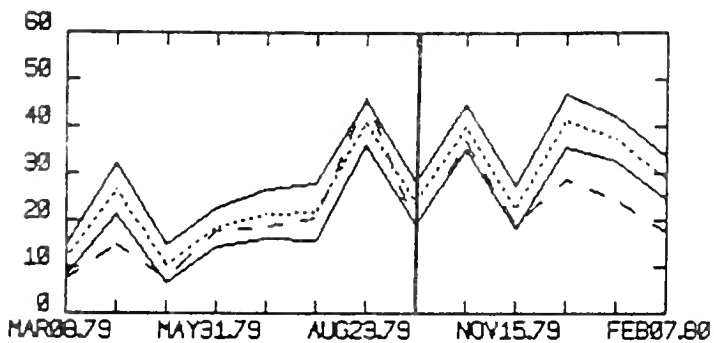
In most cases the share changes are followed within a standard error or so but sometimes the absolute level is missed. The deviation of Small B in the forecast period is an example. It would appear from the quality of fit in general that the model captures quite well the average effect of, say, promotion, but an individual promotion may be more or less successful than



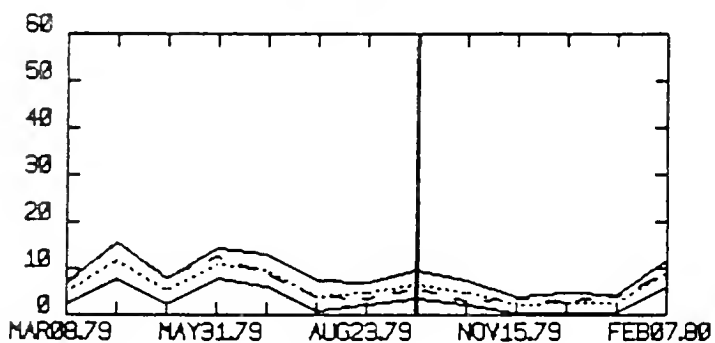
SMALL B



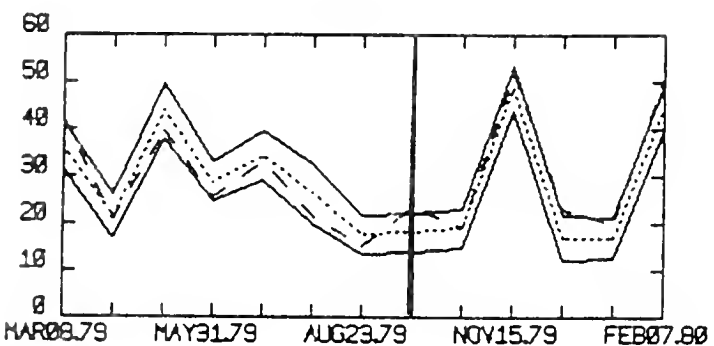
LARGE B



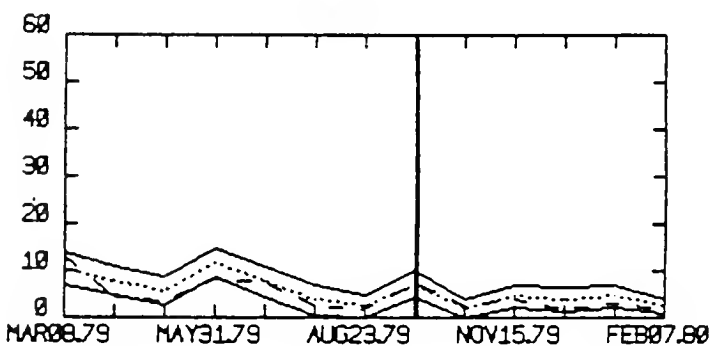
SMALL C



LARGE C



SMALL D



SMALL E

4 WEEK PERIODS

--- ACTUAL PURCHASE SHARE
 PREDICTED PURCHASE SHARE
 _____ 90% CONFIDENCE BAND

Figure 6. Tracking of purchase share in the holdout sample is quite good for most brand-sizes.

average. The one for Small B in the middle of the forecast period appears to have been less successful. The deviation persists after the promotion because the model increases the customers' loyalties on the basis of their assumed promotional purchases, whereas the actual customers experience no such effect. Despite the level difference, however, the directional changes continue to be correct.

Small D shows a level difference that puts the actual share entirely above the confidence band. Presumably, this is due to a miscalibration of the brand-size constant. On checking back, it turns out that the calibration sample contains only 31 purchases of Small D out of a total of 103 during the calibration period (3.0% share). Despite random selection of customers for the two samples, the holdout group has many more purchases of Small D, about 93 out of 916 (10.2% share) in the same period. A larger brand-size constant would probably produce a better representation of the underlying population that both samples are drawn from and would certainly improve the appearance of the Small D plot. Be that as it may, the share changes are tracked well so that, with respect to the normative questions of what happens when the decision variables are changed, the Small D plot is reassuring.

The parsimony of the model deserves emphasis. The same coefficients are used for every brand-size for all control and loyalty variables. True, there is also a set of brand-size specific constants but their principal effect is to force average predicted share to equal average actual share and this only for the calibration sample over the calibration period.

Another opportunity to test the model arises from the presence of different stores in the sample. The panel population consists of store-loyal customers from four stores, each of which has its own individual

personality. One of them is a large conventional supermarket; another a smaller, more neighborhood market; still another is a warehouse store; and the last is a conventional store with unconventional merchandising practices. The calibration process takes no account of the difference among stores beyond whatever manifestations these may generate through values of the attribute variables, nor does the calibration consider that different types of households may shop at different stores.

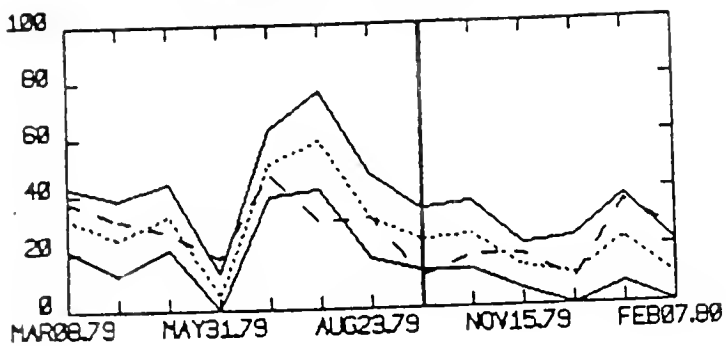
Breaking the customers down by store drastically reduces sample sizes. To compensate we aggregate over sizes to brand level and also focus on two high share brands A and C. Figure 7 presents tracking for these brands in the four stores. The format is as in Figure 5: holdout sample, predicted and actual shares with confidence band. The results are very satisfactory, with the model predicting behavior well across the four stores. Notice in particular the tracking of the decline of Brand C in Store 3. The implication is that if, we know what is happening to the control variables in the store, we can predict share quite well.

The parsimony of the model deserves emphasis. Except for the brand-size constants which, as we have discussed, determine only average share, the coefficients of the model are exactly the same for every customer, store, and brand-size. Nevertheless, they combine with the control and environmental variables to produce widely different patterns of share behavior and these correspond quite well to the actual behavior in the hold-out sample.

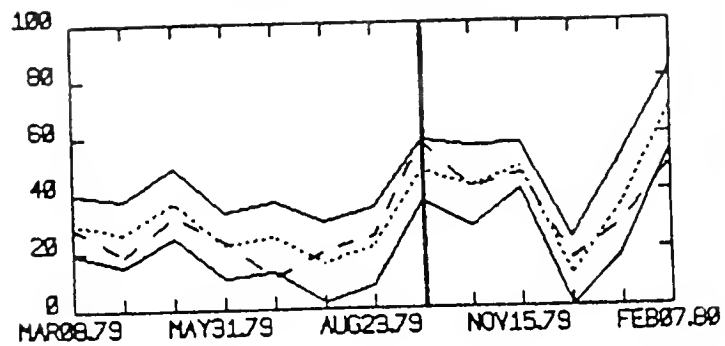
7. MARKET RESPONSE TO CONTROL VARIABLES

Practical use of the model requires measures of how the market responds to the retailers' actions.

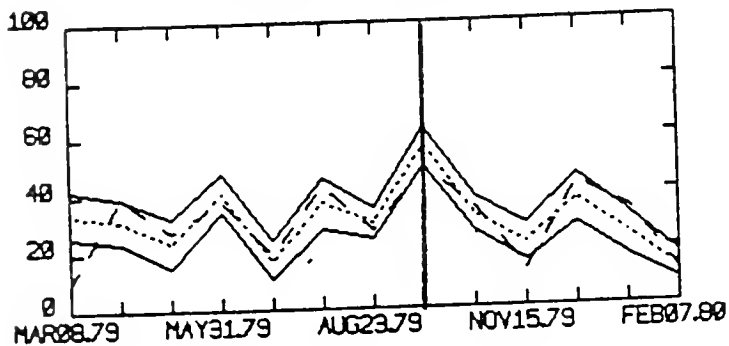
BRAND A - STORE 1



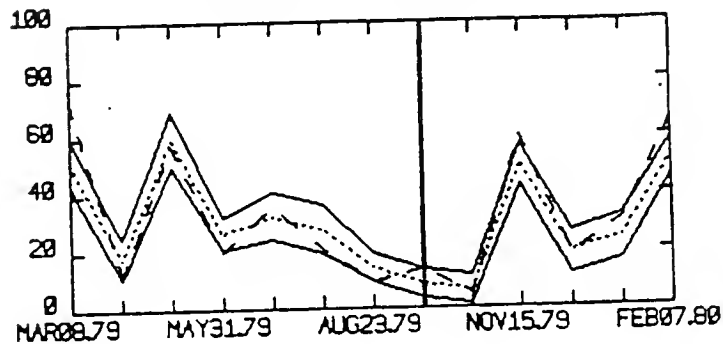
BRAND C - STORE 1



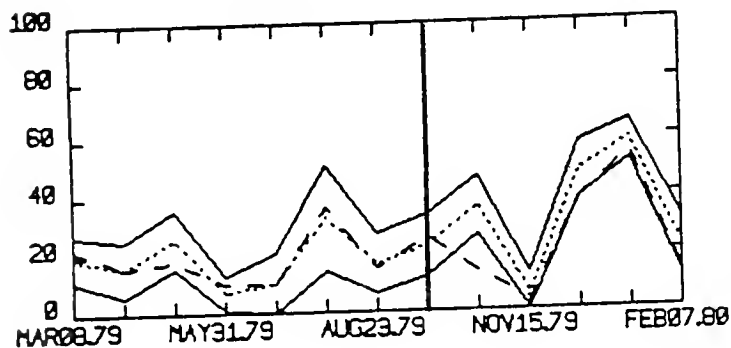
BRAND A - STORE 2



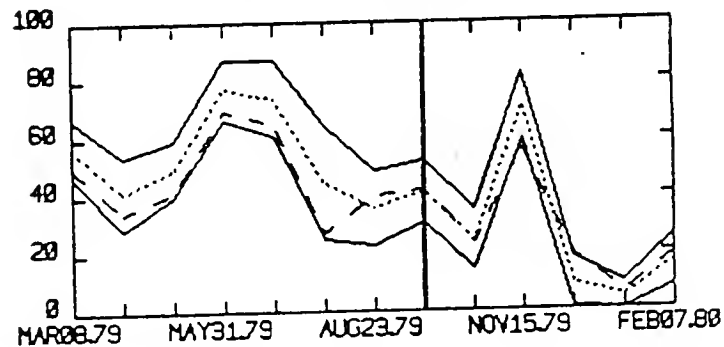
BRAND C - STORE 2



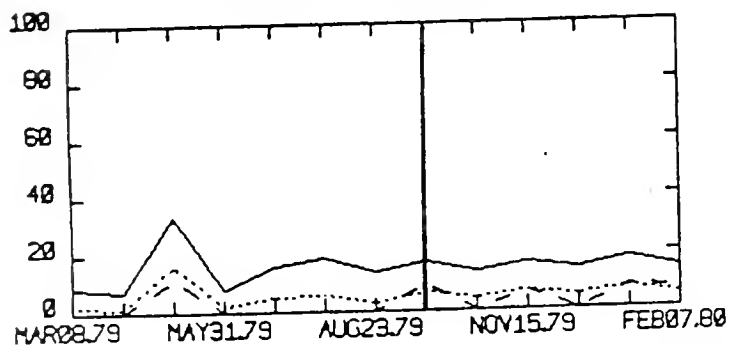
BRAND A - STORE 3



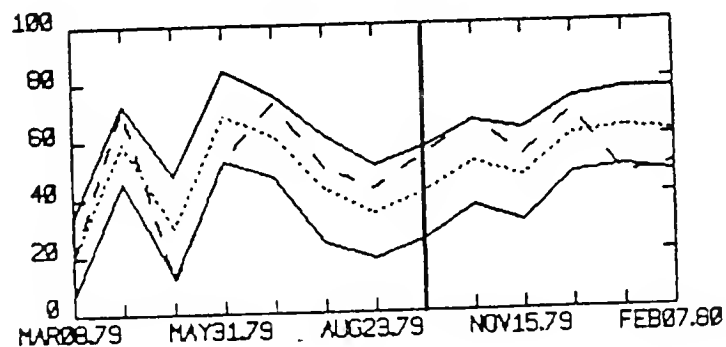
BRAND C - STORE 3



BRAND A - STORE 4



BRAND C - STORE 4



4 WEEK PERIODS

--- ACTUAL PURCHASE SHARE
 - - - PREDICTED PURCHASE SHARE
 . . . 90% CONFIDENCE BAND

Fig. 7. Tracking of purchase shares of Brands A and B in four different supermarkets is quite good despite different store characteristics.

7.1 Calculating Response

We have a calibrated model at the individual customer level but aggregate market response is usually of more interest to a decision maker and requires further computational effort. First of all, even though the logit coefficients are the same for all customers, individual responses will differ greatly depending on the store environment and prior loyalties. Table 2 illustrates this by showing that, in a hypothetical two product market, a customer with equal loyalties to both products displays a price elasticity of share equal to -2.2 , whereas a customer with 80% loyalty to one of the products has an elasticity of $-.06$ to that product. Analogously, although the control variables in our eight product market, have the same coefficients for all brand-sizes, the market response for a particular product will depend on the actions of the other brand-sizes and the distribution of loyalties across the customers.

Analytically, determination of aggregate market response calls for an integration of the customer response function over a joint distribution of customer loyalties, prices, and promotions for every brand. One standard approach to such a high dimensional integration is Monte Carlo sampling. A simpler and easier to implement method is to use the actual customers, time periods and attribute variables of the data. By changing, say, the regular shelf price of a brand-size by 1% over the whole time period and observing the corresponding change in share, we obtain an aggregate share response to shelf-prices.

Many different response calculations are possible. Some examples are (1) shelf price elasticity of share, (2) the cross elasticity of one brand-size's share to another's price, (3) share response to promotion, (4) the cannibalization of one size of a brand by the promotion of another and (5) the elasticity of price during a promotion.

HYPOTHETICAL TWO PRODUCT MARKET: A AND B

Loyalty coefficient: 7.0

Shelf-price coefficient: -30 oz/dol

No other attributes.

Price response of customer with equal loyalties

Loyalty to A = loyalty to B = .5

Price of A = price B = .15 dol/oz

Probability of choosing A = $e^{(3.5 - 4.5)} / (e^{-1} + e^{-1}) = .5$

10% price reduction of A to .135 dol/oz

Probability of choosing A = .61

Price elasticity for A = -2.2

Price response of customer with unequal loyalties

Loyalty to A = .8, loyalty to B = .2

Price of A = price of B = .15 dol/oz

Probability of choosing A = .985

10% price reduction of A to .135 dol/oz

Probability of choosing A = .991

Price elasticity for A = -.06

Table 2: Although the model coefficients are the same for all customers, response can be quite different depending on attribute values at time of purchase.

Each of these can be evaluated as a short term effect, i.e., as an alteration of choice probabilities at the purchase occasion on which the variable is charged, or as a long term effort, taking also into account the future impact of changed loyalty. The latter requires rather elaborate calculations. We have restricted ourselves to three short term examples: shelf price elasticity, response to promotion and promotional price cut elasticity. The results will illustrate a number of important points. The calculations are made as follows:

(1) Shelf-price elasticity. Using the calibration sample over the 32 week calibration period, all the shelf prices of, say, Small A are reduced by 1% and all other variables held constant. The probability of choosing Small A is calculated for all purchases to produce a Small A purchase share for the calibration period. The difference from the share without the price change is converted to an elasticity.

(2) Response to a Promotion with a Median Price Cut. Consider again Small A. For every purchase occasion calculate the probability of purchase without a promotion. Calculate the average share for Small A. Repeat the process but compute the probability of purchase when Small A has a promotion with a median price cut. All other attributes are held constant throughout. The promotion response is the percent increase of the second share over the first.

(3) Promotional Price Cut Elasticity. For every purchase occasion calculate probability of purchase with a promotion having a median price cut which is further reduced by 1% of whatever the price is. Calculate share. Take the difference from the share with promotion and median price cut and compute an elasticity.

The results are shown in Table 3.

	<u>Purchase Share (%)</u>	<u>Shelf price share Elasticity</u>	<u>Promotional share Increase (%)</u>	<u>Promotional price cut Share Elasticity</u>
Small C	28.4	-1.9	172	-1.1
Small B	17.1	-2.3	332	-1.7
Large C	15.5	-2.3	272	-1.7
Small A	14.0	-2.5	360	-1.8
Large A	10.6	-2.7	361	-1.8
Large B	9.8	-4.6	499	-2.0
Small D	3.0	-3.4	406	-3.2
Small E	1.6	-11.0	560	-2.4

Table 3: Short term market response to three different control variables.

7.2 Discussion

A first observation is that response varies widely by brand-size. This was expected even though all brand-sizes use the same coefficients for price, promotion, and promotional price cut, because the market response calculations take into account the complete environment, including competitive conditions, brand loyalties, etc. The variation in response by brand-size has significant implications for retailers and manufacturers. Differences in price elasticity, for example, suggest different pricing policies across brand-sizes.

The next point is that the response measures show a definite pattern. Table 3 has been arranged so that the brand-sizes are listed in order of decreasing share of purchases. A simple glance down the table shows a high correlation between columns. It is not perfect, thereby happily leaving some room for brand individuality, but clearly the high share brand sizes tend to have smaller elasticities. Thus, as a percentage of their own share, high share brand-sizes are less sensitive to their own control variables than low share brand-sizes. On the other hand, if we multiply the three elasticities of Table 3 by their corresponding shares we can obtain absolute share increments for the changes in control variables. If this is

done, it will be discovered that the high-share brands tend to obtain a larger share increment than the low-share brands for the same marketing action (price reduction or promotion).

Both of these phenomena are consistent with the following view of the market: each brand-size has a pool of loyal customers who are rather impervious to marketing actions, both those of their favorite brand-size and those of others. Such customers are operating on the upper flat portion of Figure 2. High share products tend to have large loyal customer pools. In addition there is a pool of switchers, who have intermediate or low loyalties across all brands. These customers are operating on the steep part of Figure 2 and respond relatively easily to marketing actions. High-share brand-sizes tend to have sold more goods to the pool of switchers than low-share brand-sizes, thereby building up an intermediate loyalty that helps increase probability of purchase for their products.

The relatively close relation between market responses and share, coming as it does from a model that predicts well, suggests the presence of a fundamental type of customer behavior in these types of markets.

As a final observation we remark on the completeness of the model. It gives information not just about a single brand-size but about the whole market. We could supplement Table 3 with other response measures as described earlier. The completeness permits the study of strategic issues, for example, deciding on the best moves to expand share, or preparing a defense against an aggressive competitor.

8. CONCLUSIONS

UPC scanner panels generate a wealth of information about market behavior, including customer response to important retailer and manufacturer

control variables. The chief advantages of the scanner panels lie in their micro detail and competitive completeness. People, not markets, respond to the actions of the retailers and manufacturers. The greater variance of the explanatory variables at the individual level offers richer opportunities for calibrating response than the corresponding store or market levels.

The multinomial logit model provides an excellent representation of the regular ground coffee shares at the retail store level. Coefficients of the model are statistically significant, many of them highly so. The calibrated model predicts the behavior of a hold-out sample of customers satisfying well. A remarkable feature of the model is its parsimony. The major coefficients, namely, those for brand and size loyalty and the control variables, are the same across all brand-sizes and all customers, and yet the model ably predicts brand-sizes with widely different shares, follows different trends and turns over time and tracks brand performance in different types of stores. The results are not perfect and, reasonably so, since various marketing phenomena are missing (for example, display quality, couponing, and media advertising) and we know from other sources that these actions influence purchases. Nevertheless, the results here seem very promising.

Manipulation of the calibrated model yields share response to several marketing variables. Here we have only scratched the surface. However, it is clear that, because we are modeling the actions of all brand-sizes, the answers produced by the model will depend strongly on the questions it is asked. In particular the answers will depend on the customer loyalties and competitive actions assumed.

Much work remains to be done. A major missing feature from a practical point of view is the modeling of the purchase occasion itself. Our work

here has focussed entirely on share, whereas some of the market actions involved, notably promotion, tend to shift purchases in time and therefore at least temporarily expand the market.

The model has further practical deficiencies. As developed here it describes what happens inside the retail store. This is not adequate for the retailer, who is relatively unconcerned about competition among brands, except perhaps for house brands, but is very concerned about competition from other retailers. For a discussion see Little and Shapiro (1980). Correspondingly, although manufacturers are strongly interested in interbrand competition, the control variables of our model are those of the retailer and only indirectly influenced by the manufacturer. Various steps can be taken to extend the results closer to the decision makers' needs but such work remains in the future.

A good representation of the coffee market is encouraging, but coffee is just one of the many categories of products in the supermarket. We do not know whether the strong assumptions of the multinomial logit will hold up in more complicated situations where, for example, product differentiation is greater or variety seeking is common or the customer maintains a portfolio of products. We expect that a variety of new modeling issues will surface and require ingenuity to resolve.

Nevertheless, in the coffee case we see the suggestion of a new understanding of the interplay between brand franchise and marketing actions. Here is a parsimonious model that fits the data well and which reveals a distinct pattern between share of purchases and market response. Since the correlations are not unity nor is tracking perfect, there appear to be idiosyncratic effectiveness variables by brand and occasion and, of course, other marketing variables await inclusion in the model.

Nevertheless at the level of general understanding of these markets, the picture emerges that a well-entrenched brand-size has a set of loyal customers who, at least in the short run, make its share relatively insensitive to certain marketing actions. At the same time the market contains another group of customers less loyal to any brand-size who switch with changes in marketing actions. More important, the exact interplay and its ramifications for brand strategy now seem to be within striking distance of our marketing science and technology.

ACKNOWLEDGEMENT

The authors gratefully thank Selling Areas Marketing, Inc. (SAMI) for providing the scanner panel data on which this work is based. We are expecially appreciative of the help of Mr. Kenneth Silvers, Mr. Frank Smith, Ms. Valerie Gager and Mr. Jack Moffly.

APPENDIX

TEST FOR INDEPENDENCE FROM IRRELEVANT ALTERNATIVES

The multinomial logit model implicitly assumes "independence from irrelevant alternatives" (IIA). McFadden, Train and Tye (1977) present a residuals test that we shall apply to evaluate whether the assumption reasonably holds in our set of data. The motivating idea of the test is that violation of the IIA property will cause systematic errors in the choice probabilities. The procedure first calculates the probabilities by the calibrated model for, say, alternative j for each of the 1021 observations. The probabilities are then ranked and sorted into some number of cells, roughly the same number of observations in each cell. For each cell we calculate an expected number of choices of j from the probabilities and compare it with the actual number in a goodness of fit test. The statistic

$$\chi^2 = \sum_{m=1}^M (S_m - N_m \bar{P}_{jm})^2 / N_m P_j$$

where m = index of cell

M = total number of cells

S_m = number of actual choices of j in cell

N_m = total number of observations in cell

\bar{P}_{jm} = average probability for alternative j in cell m

\bar{P}_j average probability for alternative j in total sample,

has an asymptotic distribution bounded by χ^2 distributions with $M-1$ and $M-K-1$ degrees of freedom where K is the number of estimated parameters. The test statistics are not independent across alternatives.

The test was run for each brand-size, dividing the observations into 50 cells each time or about 20 observations per cell. The resulting χ^2 were

	χ^2		χ^2
Small A	24.6	Small C	17.9
Large A	18.5	Large C	18.6
Small B	33.0	Small D	9.5
Large B	17.0	Small E	30.2

The degrees of freedom for the upper bound χ^2 statistic is 49 with critical (.05) level 66.1 and for the lower bound 35 and 49.7. As may be seen the χ^2 for each brand size falls well below both critical levels. Therefore no departure from IIA is detected.

REFERENCES

- Ben-Akiva, M. (1973), "Program for maximum likelihood estimation of the multinomial logit model," Transportation Systems Division, Department of Civil Engineering, M.I.T., (Jan).
- Dodson, J.A., A.M. Tybout and B. Sternthal (1978), "Impact of deals and deal retraction on brand switching," Journal of Marketing Research, 15 (Feb), 72-81.
- Domencich, T.A. and D. McFadden (1975), Urban Travel Demand, North Holland, Amsterdam.
- Gensch, D.H. and W.W. Recker (1979), "The multinomial, multiattribute logit choice model," Journal of Marketing Research, 16 (Feb), 124-132.
- Hauser, J.R. (1978), "Testing the accuracy, usefulness, and significance of probabilistic choice models: an information theoretic approach," Operations Research, 26 (May), 406-421.
- Hlavac, T.E., Jr. and J.D.C. Little (1966), "A geographic model of an urban automobile market," in D.B. Hertz and J. Melese, ed., Proceedings of the Fourth International Conference on Operational Research, Wiley-Interscience, New York.
- Jones, J.M. and F.S. Zufryden (1980), "Adding explanatory variables to a consumer purchase behavior model: an exploratory study," Journal of Marketing Research, 17 (Aug), 323-334.
- _____ (1981), "Relating deal purchases and consumer characteristics to repeat purchase probability," Journal of the Market Research Society, 23, 84-99.
- Little, J.D.C. and J.F. Shapiro (1980), "A theory for pricing nonfeatured products in supermarkets," Journal of Business, 53 (July, Part 2) S199-209.
- Luce, R.D. (1959), Individual Choice Behavior, John Wiley and Sons, New York.
- McFadden, D. (1974), "Conditional logit analysis of qualitative choice behavior," in P. Zarembka Frontiers in Econometrics, Academic Press, New York.
- _____ (1981), "Econometric models of probabilistic choice," in C.F. Manski and D. McFadden Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge.
- _____, K. Train and W.B. Tye (1977), "An application of diagnostic tests for the independence of irrelevant alternatives property of the multinomial logit model" Transportation Research Record, 637, 39-45.

- Punj, G.N. and R. Staelin (1978), "The choice process for graduate business schools," Journal of Marketing Research, 15 (Nov), 588-598.
- Shoemaker, R.W. and F.R. Shoaf (1977), "Repeat rates of deal purchases," Journal of Advertising Research, 17 (April), 47-53.
- Silk, A.J. and G.L. Urban (1978), "A pre-test-market evaluation of new packaged goods: a model and measurement methodology," Journal of Marketing Research, 15 (May), 171-191.
- Theil, H. (1969), "A multinomial extension of the linear logit model," International Economic Review, 10 (October), 251-259.
- Urban, G.L., P.L. Johnson and R.H. Brudnick (1981), "Market entry strategy formulation: a hierarchial modeling and consumer measurement approach," Sloan School Working Paper 1103-80, M.I.T. (Jan).

SEP 24 1984

01 11

BASEMENT

Date Due

APR 11 1990	
JY 21 '90	
JAN 02 1991	
MAY 03 1991	
APR 19 1991	
MAY 10 1991	
SEP 0 1991	
JUN 0 1 1998	

Lib-26-67

H028.M414 no.1304- 82
Guadagni, Pete/A logit model of brand
744308 0xBKS 00134104



3 9080 002 013 644

